Client Side Channel State Information Estimation for MIMO Communication

Sambhavi Tiwari
Department of IT, IIITA
Prayagraj, UP, India
rsi2018503@iiita.ac.in

Abhishek
Department of IT, IIITA
Prayagraj, UP, India rsi2016006@iiita.ac.in

Shekhar Verma
Department of IT, IIITA
Prayagraj, UP, India sverma@iiita.ac.in

K.P Singh
Department of IT, IIITA
Prayagraj, UP, India
kpsingh@iiita.ac.in

M. Syafrullah
Program of Master of Computer Science
Universitas Budi Luhur
Indonesia
mohammad.syafrullah@budiluhur.ac.id

Krisna Adiyarta
Program of Master of Computer Science
Universitas Budi Luhur
Indonesia
krisna.adiyarta@gmail.com

***Abstract*—Multiple-input multiple-output (MIMO) system relies on a feedback signal which holds channel state information (CSI) from receiver to the transmitter to do pre-coding for achieving better performance. However, sending CSI feedback at each time stamp for long duration is an overhead in the communication system. We introduce a deep reinforcement learning based channel estimation at receiver end for single user MIMO communication without CSI feedback. In this paper we propose to train the receiver with known pilot signals to analyse the stochastic behaviour of the wireless channel. The simulation on MIMO channel with additive white Gaussian noise (AWGN) shows that our proposed method can learn the different characteristics affecting the channel with limited number of pilot signals. Extensive experiments show that the proposed method was able to outperform the existing state-of-the-art end to end reinforcement learning method. The results demonstrate that the proposed method learns and predicts the stochastic time varying channel characteristic accurately at receiver's end.**

Keywords—MIMO, AWGN, Pilot Signal, Deep Reinforcement Learning, CSI, DQN.

I. INTRODUCTION

MIMO is being employed in various wireless systems [1] like 4G cellular and wireless LAN to increase bandwidth and coverage using multi path channel characteristics. By encoding information over multiple antenna elements using spatial diversity or spatial multiplexing schemes, the throughput and range can be improved under various channel conditions. Schemes based on hard analytical methods of coding/beam forming and decoding for these problems are found to be sub optimal. In addition, spatial multiplexing methods typically rely on the accurate estimation of CSI, quantization, and feedback, which complicates the ability of these schemes to perform optimally. Machine learning techniques are being employed to improve the performance through data driven approaches in wireless communication systems [2] [3] especially through CSI estimation. Fig.1 shows a MIMO communication system where, x is vector of the input symbols $(x \in C^{M_T})$, H is the channel matrix $(H \in C^{M_R \times M_T})$ and $w$is the vector of AWGN with receiving antennas $(w \in n^{M_R})$. Over a MIMO channel with $M_T$antennas at the transmitter and $M_R$antennas at the receiver, received symbols can be expressed as:

$$\hat{x} = Hx + w \,, \tag{1}$$

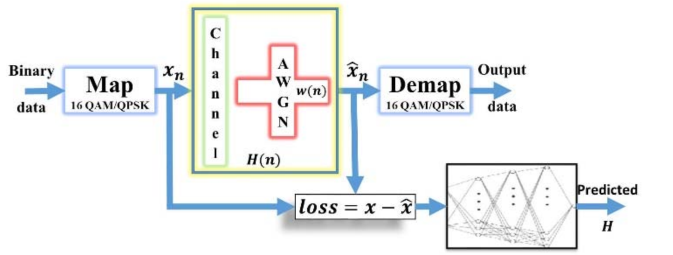

Fig. 1: Architecture of the proposed MIMO communication system with deep Q Network (DQN) for channel estimation.

There are significant results shown in the past literature on the channel parameter estimation. Authors in articles [4]– [7] used optimal linear algorithm and auto regressive tracing algorithm for prediction of channel fading, where channel impulse response prediction has been done by combining the current CSI with the past one. A support vector machine-based method is proposed to predict a more practical environment [8]. Due to large number of tuneable parameters, these models become highly complex and fail to minimize the error rate. Further, there is a need to increase the throughput of channel capacity in available bandwidth for MIMO system and one of the reasons of this decrease in throughput is due to the continuous feedback sent to transmitter. To overcome this issue in MIMO, feedback reduction methods [9] and deep neural networks are already being used.

In MIMO communication, channel knowledge is prerequisite for low error recovery of signals at the receiver. Hence, precise CSI estimation is important. In a time division duplex system, the reciprocity between uplink and downlink can be utilized wherein, the base station with enough computational capacity can estimate the CSI based on pilot signals sent by a transmitter. A pilot signal is a single frequency signal, transmitted over a communications system which is known at both transmitter and receiver end and can be defined as:

$$P_{i,j(n)} = H_{i,j(n)} + e_{i,j(n)}, \tag{2}$$

Where, $e_{i,j(n)}$is the error estimation of the channel. The CSI is used for appropriate link adaptation for receiving signals from the mobile stations. As the mobile stations are resource constrained, this estimated uplink CSI can be broadcast by the base station to the mobile stations. The mobile stations can utilize this CSI feedback for precoding, for uplink or as CSI for the downlink. However, in a time varying channel, the periodic feedback with associated delay and extra signalling overhead have negative impacts on system performance. Moreover, if uplink and downlink channel characteristics are not reciprocal, like in frequency division duplex, signal recovery is adversely affected. It is necessary to estimate the downlink CSI on mobile stations for low error signal recovery. This requires one, lightweight CSI estimation and two, a prediction mechanism for CSI to avoid frequent periodic transmission of downlink pilot signals.

The problem is to design a CSI estimator which can learn H with minimal pilot signals for low communication overhead. However, the channel matrix is time varying. Hence, there should be a mechanism that can predict this time varying channel state accurately. Additionally, the overall mechanism should be computationally light for possible implementation on mobile receivers. In this paper, we propose a lightweight deep reinforcement learning (DRL) based channel estimator for MIMO systems at receiver end. The proposed model exploits the properties of known pilot signal to quickly learn the dynamic behaviour of the channel. The accurate predictions for the next time instances allow the receiver to extract the desired signal optimally.

The major contributions of the proposed work are listed as follows:

- The lightweight Q-learning approach with three hidden layers solve the Markov Decision Process (MDP) of estimating the next CSI exploits the channel characteristics specific to each mobile equipment.
- It increases the throughput of MIMO channels by removing the CSI feedback overhead.
- By training the individual mobile receiver's using very limited number of pilot signals, obsoletes the requirement of heavy channel estimation computation at the base station.

II. DQN FOR MIMO

A. Channel state information(CSI):

Wireless CSI is the known channel properties of the communication link. This information helps in describing how a signal propagate over the wireless channel. CSI is estimated at receiver's end and sent to the transmitter through feedback link. In order to increase the transmission rate, coverage, spectral efficiency and to reduce receiver complexity, CSI is used at transmitter [10]. MIMO relies on CSI feedback for pre-coding and due to the dynamic nature of the channel, CSI at the transmitter is gaining more importance in MIMO [11]. In cellular scenario, getting accurate CSI feedback at a transmitter faces the following challenges:

- The inherent delay caused by processing time at the receiver end.
- Feedback overhead containing pre-coding, channel quality, and rank indicator.
- The number of feedbacks increases with the increase in number of terminals.
- Due to high mobility of mobile devices/receivers the transmitter receives infrequent/limited CSI feedback, it exhibits link adaptation based on short-term feedbacks.

However, many state-of-the-art methods are proposed to overcome these above challenges. In [12] the study mainly focused on reducing feedback overhead by using the spatial and temporal correlation of CSI. Several techniques for improving CSI in connection with full dimension-MIMO have been proposed [13]. Further studies show predicting CSI by the linear Wiener predictor if CSI is aged [14].

B. RL Strategy

This section explains the use of deep Q network and deep reinforcement learning-based strategy in MIMO communication system to predict the dynamic behaviour of the channel. It has been seen that for high-dimensional data, deep Q network learns the control strategies directly. The input fed into DQN is the states, or observations that finally return the value function corresponding to each action as output. Deep Q network is a combination of *Q learning* and *Deep neural network*. Q learning is a method of solving sequential decision problems by learning estimates of each action by its optimal value called as expected total sum of future rewards received by taking the most optimum action which follows an optimal policy. For a given policy π, the true value for an action $a$in state s is:

$$Q_{\pi}(s,a) = \mathbb{E}[R_1 + \gamma R_2 + \ldots | S_0 = s, A_0 = a, \pi], \qquad (3)$$

Where,$\gamma \in [0,1]$is called as discount factor which trades off the importance of immediate reward and later rewards. The Q learning algorithm is basically computing two things: *An optimal value*, $Q_*(s,a) = max_{\pi} Q_{\pi}(s,a)$.

An optimal policy calculated from the optimal values by selecting the highest valued action for each state. Estimates for the optimal action values can be learned using Q-learning [15], a form of temporal difference learning [16]. High dimensional data processing tasks are increasing everyday which cannot be solved by mere Q updates in real time. It is not possible to learn all action values in every state separately. For solving such real complex problems deep Q network is used. In DQN, we can learn a parameterized value function $Q(s,a;\theta_t)$. In the proposed work, DQN is implemented at receiver side of MIMO communication system which uses pilot signals to predict the stochastic channel behaviour without using feedback.

III. EXPERIMENT AND RESULTS

In this section, we have constructed a deep reinforcement model which can learn the behavior of 2X2 MIMO communication channel, where we used statistical performance-based evaluation of deep Q network using pilot signals.

A. MIMO setup:

The proposed deep learning prediction model of a deep Q network was able to predict the channel information for different time stamp t in 2X2 MIMO communication system. Our work has considered 2X2 MIMO communication system in which a transmitter sends data using 2 transmit antennas (M_T=2), 4 channels (H_n=4), and 2 receiver antennas(M_R=2).

B. Deep-RL setup for MIMO:

In this proposed work, following entities represent the MIMO system as a deep Q learning architecture:

- An *environment* which consists of transmitter (M_T), channel (H_n), receiver (M_R)and pilot signals (P_n).
- *States* are the pilot signals at each time stamp i.e., $P_n = x_n$
- An *agent* is the channel.
- *Actions* are the transmitted signals from the channel $\hat{x}_n$.
- *Utility* value that maps difference between x and $\hat{x}$. *Utility* here is inversely proportional to *reward*. Therefore, we consider *reward* value as (1−*utility*). Fig 1 shows all elements of our proposed architecture.
- *Policy*, our model need exploration. Hence, we used ϵ greedy policy i.e., it is a way of selecting random actions with uniform distribution from a set of available actions.

For the performance evaluation, our proposed training scheme used deep Q learning at receiver's end. Further, compared with training scheme used in existing method [17] using feedback and training done at both receiver and transmitter ends. Evaluations are done on AWGN channel. However, channel behaves in stochastic manner, whose output $\hat{x}$ follows conditional probability distribution on input$x$i.e.,$\hat{x} \sim P(y|x)$.

Further, Deep Q network learning steps are shown in algorithm 1. All the processing and prediction of the signals are done at *receiver side*. Therefore, initially in our system we knew the pilot signals P_i, and channel H_i. The channel H_i was calculated from known pilot and received signal.To simulate time-varying MIMO channels for generating training samples we have created simulation environment using python. For 2*X*2 MIMO, M_R = 2, M_T= 2, H = 4 with AWGN (μ = 0.0, σ = (0.1,0.2,0.3…,0.9)). The simulation results are computed for 1000 iterations for 10 times. The learning rate was set 0.01 for all iterations and the simulation code was run for 9 different σ values i.e., (0.1, 0.2, 0.3... 0.9), to learn the dynamic behavior of the channel as shown in TABLE I.

Algorithm 1: DQN for MIMO

Input: Pilot signal $\boldsymbol{P_i = \hat{x}_i}$**,** $\boldsymbol{H_i}$

Output: Predicted H i.e., $\hat{H}$. Initialize action value function Q with random weights $\boldsymbol{\theta}$.

Let $\boldsymbol{H'(H_1,H_2,H_3,\ldots,H_i)}$be available channel characteristics with loss $\boldsymbol{L}$at receiver's end of MIMO communication system. For time stamp $\boldsymbol{i}$ to $\boldsymbol{n}$, where $\boldsymbol{i = (1,2,3,4,..,n)}$ **do:**

1. Compute the reward r between P_i and received signal $\hat{x}$ at *receiver end*.

$$Loss = (P_i - \hat{x}_i) \quad (4)$$

$$r = (1 - r) \quad (5)$$

2. With probability Ɛ = 0.2
   Choose random H as $\hat{H}$.
   Select $\hat{H} = max_H Q(P_i, H_i; \theta)$
3. Set $\hat{H}_i$=

$$\begin{cases} r_i & ; if\ (i = n) \\ r_i + \gamma max_{H'} Q\big(P_{i+1}, H'; \theta\big); & otherwise \end{cases}$$

4. Perform a gradient descent step on $(\boldsymbol{H_i} - \boldsymbol{Q(P_i, H_i; \theta)})^2$ with respect to the network parameters $\boldsymbol{\theta}$.
5. Stop when**, i = n**.

End for

TABLE I: Minimum, maximum, mean, standard deviation values of predicted $\hat{H}$ using different noise (AWGN) for 1000X10 iterations.

Noise	Minimum value	Maximum Value	Mean(μ)	SD
σ=0.1	0.009591	0.246794735	0.042042842	0.031069587
σ=0.2	0.000442263	0.905297033	0.107660084	0.087653616
σ=0.3	0.000420945	0.845623376	0.153214927	0.153214927
σ=0.4	0.000364149	1.495377705	0.226484045	0.172162198
σ=0.5	0.00056568	2.284914765	0.313134364	0.247266636
σ=0.6	0.000280382	2.772616411	0.407752488	0.326373243
σ=0.7	0.000792498	4.285453473	0.508756603	0.41437401
σ=0.8	0.000239953	4.557037772	0.621707027	0.502095435
σ=0.9	0.000822981	6.512823511	0.763946241	0.638754053

For the data transfer over 2X2 MIMO communication system case, shown in graph (Fig.2), we can see that our method was able to predict the channel characteristics. In this graph blue line represents actual H values i.e., $\hat{H}_{MSE}$ and red line in graph represents predicted $\hat{H}$ i.e. $\hat{H}_{MSE}$. The graph shows comparison between H *and* $\hat{H}$ for initial 100-time stamps having mean squared error values. It is clearly visible in the graph, maximum 5 known pilot signals (*P(i=1, 2…, 5)*) are sent to predict approximately next 10 time slot change $\hat{H}_{(i=1,2,3,\ldots,10)}$ in the channel behavior. For an instance we can see that after t = 3 deep Q network predicts the next H with minimum loss and continues its prediction till t = 39-time stamp, further again pilot signals are sent for next 2-time stamps (t = 41 − 42). Afterwards, our model again starts learning the stochastic nature of the channel to predict for next time stamps with minimum loss. Reinforcement learning based prediction can be seen clearly in TABLE II.

TABLE II: Estimation of total number of counts for sending pilot signals and H prediction for future time slots.

Pilot signals sent for t time slots	Predicted $\hat{H}$ for next time slots
1-3	14
15-16	17-27
28-30	31-42
43-44	45-54
55-58	59-66
67-68	69-74

75-76	78-90
91-93	94-100

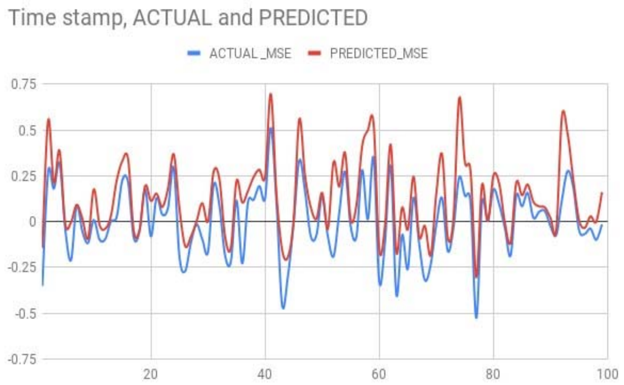

Fig.2. Smooth line plot of original channel generated by 2X2 MIMO communication model vs predicted channel using our method.

C. Evaluation results and comparison

Fig. 3 shows the comparison result for σ = 0.3 of proposed algorithm prediction scheme with that of *End-to-End Learning of Communication Systems without a Channel Model* [17]. Though in this approach algorithm iterates between supervised training of receiver still the output accuracy of the model is not efficient as our proposed method. Our method works on real time training of the systems and was able achieve better performance and predicting nearly close to actual values of simulated *H* values.

- In contrast with *End-to-End Learning of Communication Systems without a Channel Model* [17], the proposed method remains light weight, the same has been stated in table (TABLE 3).
- The TABLE.3 compares the execution time of both algorithms on a standard personal computer with the following configurations:
  Processor: Intel(R) Core (TM) i7-8700CPU @ 3.20GHz 3.19GHz
  RAM: 8.00 GB
  System type: 64-bit Operating System, x64 based processor.
- As in the proposed system, only receiver M_R is trained *online* with minimal pilot signal streams as shown in TABLE III, the lightweight architecture of Q-update mechanism with a smaller number of parameters/weights allow faster training. Whereas, in existing state-of-the-art method, both M_T and M_R are trained by averaging the effect on each channel, but spatially mobile devices may get affected from different channel characteristics. Hence, the existing method gives sub-optimal results.
- The training at each receiver M_R end by observing pilot signals individually, enables them to identify and discard the adverse effects of noise (AWGN) on channel specific to them.

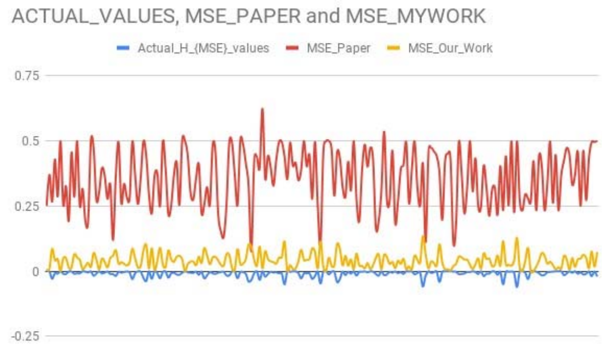

Fig.3. Smooth line graph comparison between our predicted $\hat{H}$ values, actual values of H and existing state-of-the-art H values.

- In our case, base station is no longer required for broadcasting the channel matrix. Hence, proposed method reduces the overall power consumption in MIMO.

TABLE. III: Average run time of proposed method in seconds(s).

Noise	End-to-End Learning of Communication Systems Without a Channel Model [17] (Time in seconds)	Proposed method (Time in seconds)
σ=0.1	21.061746835708618	0.5155055522918701
σ=0.2	21.880775928497314	0.5238287448883057
σ=0.3	22.33102512359619	0.5310940742492676
σ=0.4	25.664973974227905	0.5311009883880615
σ=0.5	22.02157759666443	0.5311264991760254
σ=0.6	25.878480434417725	0.5155048370361328
σ=0.7	23.455188274383545	0.5404367446899414
σ=0.8	27.42771053314209	0.5467181205749512
σ=0.9	27.72638463973999	0.5623693466186523

CONCLUSION

In this work, we proposed a novel method of estimating the future stochastic channel characteristics by sending some initial known pilot signals for MIMO using deep Q learning. The proposed scheme enhanced the channel prediction performance significantly without using CSI feedback for light weight communication systems. In contrast to the state-of-the-art method our proposed method achieved better prediction of CSI in real time for extended time slots using fewer pilot signals. This work has the future scope of extending this method on Massive-MIMO systems for other channel impairments like different types of *fading*, *scattering*, *power decay with distance*, and *attenuation*.

REFERENCES

[1] A. El-Mougy, M. Ibnkahla, G. Hattab, and W. Ejaz. Reconfigurable wireless networks. *Proceedings of the IEEE*, 103(7):1125–1158, July 2015.

[2] Yudi Huang, Junjie Tan, and Ying-Chang Liang. Wireless big data: transforming heterogeneous networks to smart networks. *Journal of Communications and Information Networks*, 2(1):19–32, Mar 2017.

[3] Alexandra Duel-Hallen, Hans Hallen, and Tung-Sheng Yang. Long range prediction and reduced feedback for mobile radio adaptive ofdm systems. *IEEE Transactions on Wireless Communications*, 5(10):2723–2733, 2006.

[4] Prachee Sharma and Kavitha Chandra. Prediction of state transitions in rayleigh fading channels. *IEEE transactions on vehicular technology*, 56(2):416–425, 2007.

[5] Alberto Arredondo, Kapil R Dandekar, and Guanghan Xu. Vector channel modeling and prediction for the improvement of downlink received power. *IEEE Transactions on Communications*, 50(7):1121–1129, 2002.

[6] Mikael Sternad and Daniel Aronsson. Channel estimation and prediction for adaptive ofdm downlinks. In *IEEE Vehicular Technology Conference*, volume 2, pages 1283– 1287. IEEE; 1999, 2003.

[7] Xiaonan Zhao, Chunping Hou, and Qing Wang. A new svm-based modeling method of cabin path loss prediction. *International Journal of Antennas and Propagation*, 2013, 2013.

[8] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews. An overview of limited feedback in wireless communication systems. *IEEE Journal on Selected Areas in Communications*, 26(8):1341–1365, October 2008.

[9] Ezio Biglieri, Robert Calderbank, Anthony Constantinides, Andrea Goldsmith, Arogyaswami Paulraj, and H Vincent Poor. *MIMO wireless communications*. Cambridge university press, 2007.

[10] Thomas L Marzetta and Bertrand M Hochwald. Fast transfer of channel state information in wireless systems. *IEEE Trans. Signal Processing*, 54(4):1268–1278, 2006.

[11] Ping-Heng Kuo, HT Kung, and Pang-An Ting. Compressive sensing based channel feedback protocols for spatiallycorrelated massive antenna arrays. In *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 492–497. IEEE, 2012.

[12] Wenting Chang, Yuan Zhu, Yushu Zhang, Huaning Niu, Gang Xiong, and Qinghua Li. Csi (channel state information) enhancements for fd (full dimension)-mimo (multiple input multiple output), February 28 2019. US Patent App.Kien T Truong and Robert W Heath Jr. Effects of channel aging in massive mimo systems. *arXiv preprint arXiv:1305.6151*, 2013.

[13] Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8(3-4):279–292, 1992.

[14] Richard S Sutton, Andrew G Barto, and Ronald J Williams. Reinforcement learning is direct adaptive optimal control. *IEEE Control Systems Magazine*, 12(2):19–22, 1992.

[15] F. A. Aoudia and J. Hoydis. End-to-end learning of communications systems without a channel model. In *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, pages 298–303, Oct 2018.